## REMARKS/ARGUMENTS

These remarks are in response to the Final Office Action dated June 15, 2004. Claims 1-23 are pending in the present application. Claims 1-23 have been rejected. Claims 1-23 remain pending. For the reasons set forth more fully below, Applicant respectfully submits that the claims as presented are allowable. Consequently, reconsideration, allowance, and passage to issue are respectfully requested.

In the event, however, that the Examiner is not persuaded by Applicant's arguments, Applicant respectfully requests that the Examiner enter the arguments to clarify issues upon appeal.

### Claim Rejections - 35 U.S.C. §103

The Examiner has stated:

> **Claims 1-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Ziauddin (U.S. Patent No. 6,029,163) in view of Schiefer et al. (U.S. Patent No. 5,761,653).**
>
> **As to Claim 1, Ziauddin teaches a method, for use in query optimization in a relational database management system (See Abstract), the method comprising the steps of:**
>
> **(a) generating statistical information regarding data which represents the results of an operation involving one or more columns of a database (See column 3, lines 5-22).**
>
> **Ziauddin does not teach (b) deriving a statistical soft constraint from the statistical information that reflects a statistical property of the data; and**
>
> **(c) using the statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan.**
>
> **Schiefer et al. teaches a method for estimating cardinalities for query processing in a relational database management system (See Abstract), in which he teaches deriving a statistical soft constraint from the statistical information that reflects a statistical property of the data (See column 8, lines 3-32, where the statistical equation "|C1|*|C2|*ff_3=100*1950*1%+1950" is read on "statistical soft constraint"; also see column 10, lines 26-13); and using the statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan (See column 2, lines 37-41; column 5, lines 63-67; column 6, lines 18-36; column 8, lines 3-27, where the statistical equation "|C1|*|C2|*ff_3=100*1950*1%+1950" is read on "statistical soft constraint"; also see column 10, lines 26-13).**

**Therefore, it would have been obvious to a person having ordinary skill in the art at the time of the invention was made to have modified Ziauddin, to include deriving a statistical soft constraint from the statistical information that reflects a statistical property of the data; and using the statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan.**

**It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Ziauddin, by the teachings of Schiefer et al. because deriving a statistical soft constraint from the statistical information that reflects a statistical property of the data; and using the statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan would help the relational database management system to use the query optimizer to analyze how to best conduct the user's query of the database with optimum speed in accessing the database (See Schiefer et al., column 1, lines 48-51). ...**

**Response to Arguments**

**In response to applicants' arguments regarding independent claims 1, 21, and 23, that neither reference teaches or suggest, "deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data, "the arguments have been fully considered but are not deemed persuasive, because Schiefer et al. teaches "unique keys" to estimate cardinalities. "Soft constraint" is read on "unique keys", which are derived using statistical functional dependencies and information contained in the catalog (See Schiefer et al., column 8, lines 3-7, lines 29-36).**

**In response to applicants' arguments regarding independent claims 1, 21, and 23, that neither reference teaches or suggest, "the statistical soft constraint of the present invention is derived from statistical information of data representing an operation involving one or more columns of a database. With the statistical soft constraint, there is no limitation as in Schiefer where no two rows can contain the same value for the columns" the arguments have been fully considered but are not deemed persuasive, because although "no tow rows can contain the same value for the columns" does not mean that "one or more columns of a database" are not used to calculate the "statistical constraint".**

**In response to applicants' arguments regarding "Ziauddin in view of Schiefer does not teach or suggest the combination of steps of the present invention", the argument has been fully considered but is not found to be persuasive, because Ziauddin teaches collecting statistics from multiple columns to represent distinct cardinality(See column 3, lines 5-11), disclosing that there is no limitation as in "soft constraints". Schiefer et al. teaches "unique keys" to estimate cardinalities. "Soft constraint" can be read on "unique keys" which are derived using statistical functional dependencies and information contained in the catalog (See column 8, lines 3-7, lines 29-36). Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Ziauddin, by the teachings of Schiefer et al. because deriving a statistical soft constraint from the statistical information that reflects a statistical property of the data; and using the statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan would help the relational database management system to use the query optimizer to analyze how to best to conduct the user's query of the database with optimum speed in accessing the database (See Schiefer et al., column 1, lines 48-51).**

**In response to applicants' argument's that independent claim 16, "Ziauddin in view of Schiefer does not teach or suggest this feature", the argument's have**

**been fully considered but it is not found to be persuasive, because Ziauddin teaches collecting statistics from multiple columns to represent distinct cardinality (See column 3, lines 5-11), disclosing that there is no limitation as in "Soft constraints". Schiefer et al. teaches "unique keys" to estimate cardinalities. "Soft constraint" can be read on "unique keys" which are derived using statistical functional dependencies and information contained in the catalog (See column 8, lines 3-7, lines 29-36). Therefore, combining Ziauddin with Schiefer et al., does disclose a "statistical soft constraint" derived from statistical information.**

Applicant respectfully disagrees with the Examiner's rejections. For the Examiner's convenience, independent claims 1, 16, 23, and 21 are reproduced in their entirety herein below.

Claims 1, 16, 21, and 23

1. (original) A method, for use in query optimization in a relational database management system, said method comprising the steps of:

(a) generating statistical information regarding data which represents the results of an operation involving one or more columns of a database;

(b) deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data; and

(c) using said statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan.

16. (original) A database management system comprising:

means for generating statistical information regarding data which represents the results of an operation involving one or more columns of a database;

means for generating a statistical soft constraint using said statistical information; and

means for utilizing said statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan.

21. (original) A computer program product comprising:

(a) a computer readable medium;

(b) code means contained in said medium for instructing a computer to perform the steps of:

(i) generating statistical information regarding data which represents the results of an operation involving one or more columns of a database;

(ii) deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data; and

(iii) using said statistical soft constraint to estimate a cardinality value for the result of a query predicate in a query plan.

23. (original) A computer readable medium containing program instructions for use in query optimization in a relational database management system, said program instructions for:

(a) generating statistical information regarding data which represents the results of an operation involving one or more columns of a database;

(b) deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data; and

(c) using said statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan.

The present invention provides a method for using statistical constraints to assist in estimating the cardinality of predicates by an optimizer of a relational database management system. Statistical information regarding data is generated where the data represents the results of an operation involving one or more columns of a database. Next, a statistical soft constraint is derived from the statistical information. The statistical soft constraint reflects a statistical property of the data. Finally, the statistical soft constraint is used to estimate a cardinality value for the result of applying one or more query predicates in a query plan. In one aspect of the present invention. (Summary.) Ziauddin in view of Schiefer does not teach or suggest these features, as discussed below.

Ziauddin discloses a method and system for collecting query workload based statistics on column groups identified by a relational database management system (RDBMS) optimizer and for identifying columns for which statistics collection is to be performed. The system collects workload statistics that are dependent on multiple columns, rather than merely single columns. Multi-column statistic generation provides more accurate results for columns having correlated data, and therefore leads to better estimated cost analysis by an RDBMS optimizer.

Schiefer discloses a method for estimating cardinalities for query processing in a relational database management system. The method is suitable for use with a query optimizer for estimating cardinalities for sets of columns or keys resulting from a grouping operation or a duplicate removal operation. (Abstract.) The method estimates cardinalities for a key formed from a grouping of columns in a table for use in a query optimizer for a relational database management system. Selectivities and keys associated with columns in the table are provided in a catalog. The method includes the steps of: (a) determining an equivalence class for each

column in the key; (b) for each equivalence class, determining an effective cardinality for each of the columns belonging to the equivalence class; (c) determining a cardinality for each of the equivalence classes by choosing the minimum effective cardinality for the columns belonging to the equivalence class; and (d) estimating a cardinality value for the key from the product of said cardinalities for the equivalence classes. (Column 3, lines 46-60.)

Ziauddin in view of Schiefer does not teach or suggest, "deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data," as recited in independent claims 1, 21, and 23. Applicants agree with the Examiner that Ziauddin does not teach deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data.

Applicants respectfully assert that Schiefer also does not teach or suggest, "deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data," as recited in independent claims 1, 21, and 23. The Examiner has referred to "unique keys" as being the same as the statistical soft constraint of the present invention. However, Schiefer does not teach that a unique key is a "statistical soft restraint" as recited in the present invention. Instead, Schiefer teaches that a unique key is a column or group of columns. Schiefer specifically states that a "column is a statistically unique key if the cardinality of the column is very close to the cardinality of the table" (column 8, lines 3-5). Schiefer further states that a "unique key is formed from a set of columns in a table where no two rows contain the same value for these columns" (column 8, lines 28-47). Schiefer further states that "statistically unique keys comprise a column (or set of columns) which has a cardinality that is very close, i.e. 95%, to the cardinality of the table" (column 9, lines 3-5). In accordance with the present invention, the

statistical soft constraint is not a column or group of columns but is instead derived from statistical information of data representing an operation involving one or more columns.

The Examiner has also stated that a soft constraint reads on unique keys, "which are derived using statistical functional dependencies and information contained in the catalog," reciting column 8, lines 3-7 and lines 29-36. However, as described above, Schiefer specifically states that a "unique key is formed from a set of columns in a table where no two rows contain the same value for these columns" (column 8, lines 28-47). Furthermore, Schiefer specifically defines a functional dependency as follows: "column A functionally determines column B, if for any two rows the columns agree on the value for A, then they also agree on the value for B" (column 7, lines 33-37). A unique key and a functional dependency are clearly different concepts (column 6, lines 31-36 and lines 46-48) and both are clearly different from the statistical soft constraint of the present invention, which is "derived from statistical information of data representing an operation involving one or more columns."

Applicants respectfully assert that Schiefer also does not teach or suggest "using said statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan," as recited in independent claims 1, 21, and 23. Applicants agree with the Examiner that Ziauddin also does not teach this feature. As described above, the "unique key" is not a statistical soft constraint but is instead a column or group of columns.

Since Schiefer does not teach or suggest "deriving a statistical soft constraint from said statistical information that reflects a statistical property of said data," or "using said statistical soft constraint to estimate a cardinality value for the result of applying one or more query predicates in a query plan," as recited in independent claims 1, 21, and 23, Schiefer cannot be combined with Ziauddin to provide the present invention. Therefore, Ziauddin in view of

Schiefer does not teach or suggest the combination of steps of the present invention, and these claims are allowable over Ziauddin in view of Schiefer.

Independent claim 16

Independent claim 16 recites a "statistical soft constraint using said statistical information." As described above, with respect to independents claims 1, 21, and 23, Ziauddin in view of Schiefer does not teach or suggest this feature. Accordingly, the above-articulated arguments related to independent claims 1, 21, and 23 apply with equal force to independent claim 16. Therefore, claim 16 is allowable over Ziauddin in view of Schiefer for at least the same reasons as claims 1, 21, and 23.

Remaining dependent claims

Dependent claims 2-15 and 17-20 depend from independent claims 1 and 16, respectively. Accordingly, the above-articulated arguments related to independent claims 1 and 16 apply with equal force to claims 2-15 and 17-20, which are thus allowable over the cited references for at least the same reasons as claims 1 and 16.

Conclusion

In view of the foregoing, Applicant submits that claims 1-23 are patentable over the cited references. Applicant, therefore, respectfully requests reconsideration and allowance of the claims as now presented.

Applicant's attorney believes that this application is in condition for allowance. Should any unresolved issues remain, the Examiner is invited to call Applicant's attorney at the telephone number indicated below.

Respectfully submitted,
SAWYER LAW GROUP LLP

August 16, 2004
Date

Joseph A. Sawyer, Jr.
Attorney for Applicant(s)
Reg. No. 30,801
(650) 493-4540